Delivery 3

Daniel Salgado Rojo

Data Visualization and Modelling. Course 2017/2018.

9) We have two sensors that are meant to detect extreme temperature, which occurs 20% of the time. The sensors have identical specications with a false positive rate of 1% and a false negative rate of 3%. If the power is o (dead battery), the sensors will read negative regardless of the temperature. Suppose now that we have two sensor kits: Kit A where both sensors receive power from the same battery, and Kit B where they receive power from independent batteries. Assuming that each battery has a 0.9 probability of power availability, compute the probability of extreme temperature given each of the following scenarios:

   (a) The two sensors read negative.
   (b) The two sensors read positive.
   (c) One sensor reads positive while the other reads negative.

   Answer the previous questions with respect to each of the two kits. You must use direct computation "by hand" as well as `gRain`.

   In which of the three cases the answer is the same for the two kits? Why?

   Solution. Given the information of the problem and the causality dependences on the variables: Temperature, Battery (or batteries) and the two Sensors,

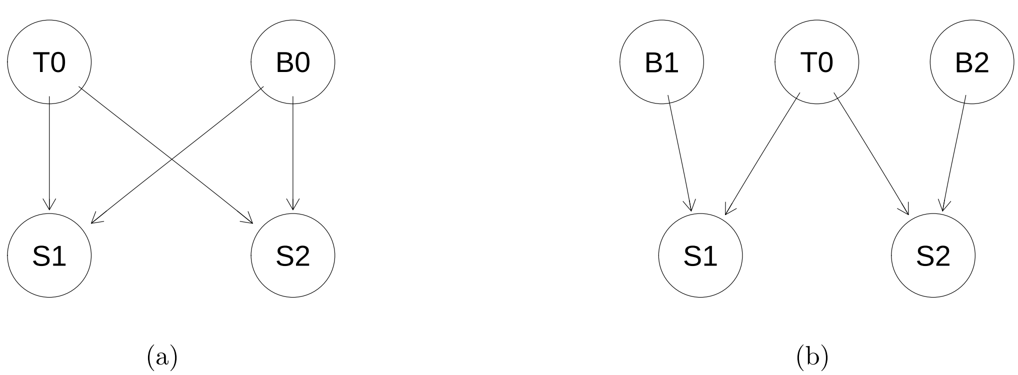

Figure 1: DAG of the Bayesian Network representing (a) kit A (b) kit B.

we deduce that the DAG corresponding to the Bayesian network that models the situations of kit A and kit B correspond to those in Figure 1.

If we have a complete set of conditional probability tables corresponding to the DAGs above, then applying Theorem 2 from the slides we will able to conclude that we have in fact a Bayesian network, one for each kit of sensors.

For the variable temperature T_0 which is true if there is extreme temperature and false if not, has the marginal from Table 1

Table 1: Marginal probabilities of T_0

T_0	$P(T_0)$
T	0.2
F	0.8

Equivalently all the sensors B_i (for $i = 0$ if we consider kit A and $i = 1, 2$ if we consider kit B) have the marginal probabilities from Table 2. Where true means that the battery is working and false means that it is not (i.e., off).

Table 2: Marginal probabilities of the batteries B_i.

B_i	$P(B_i)$
T	0.9
F	0.1

Finally, the conditional probability table for the sensors S_1 and S_2 are the ones from Table 3, Where $f_p = 0.01$ and $f_n = 0.03$. And for the kit A B_i is the same battery B_0 for both S_1 and S_2; and for the kit B, S_1 has associated the battery B_1 and S_2 has associated the battery B_2 as it is represented in Figure 1.

Table 3: Conditional probability table for the sensors with respect to its parents.

T_0	B_i	S_i	$P(S_i/T_0, B_i)$
F	F	F	1
F	F	T	0
F	T	F	$1 - f_p$
F	T	T	f_p
T	F	F	1
T	F	T	0
T	T	F	f_n
T	T	T	$1 - f_n$

Translating the probabilities we are asked for to the notation we have introduced above we have that:

(a) The two sensors read negative is equivalent to find the probability

$$P(T_0 = T/S_1 = F, S_2 = F),$$

(b) The two sensors read positive is equivalent to find the probability

$$P(T_0 = T/S_1 = T, S_2 = T),$$

(c) One sensor reads positive while the other reads negative is equivalent to find the probability

$$P(T_0 = T/S_1 = T, S_2 = F) = P(T_0 = T/S_1 = F, S_2 = T).$$

As a convention of notation $\{T_0\}$ will be equivalent to $\{T_0 = T\}$ (so if we no specify the value it is assumed to be "T"). Moreover, X_F and X_T will denote the events $\{X = F\}$ and $\{X = T\}$ respectively.

In order to not being to dense in the "by hand" resolution we don't specify always which property are we using to go from one equality to the other. Results and properties that we are going to use given the fact that we work with the Bayesian networks with the DAGs from Figure 1 and the conditional probability tables from tables 1, 2 and 3:

- The law of total probability,
- Bayes' Theorem for the conditional probabilities,
- That a node is conditionally independent from any non descendent given its parents, etc.

1 Solutions by hand

1.1 Kit A

Case (c)

$$P(T_0 = T/S_1 = F, S_2 = T) = \frac{P(S_1 = F, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = T)}$$

- Numerator term,

$$\begin{aligned} P(S_{1F}, S_{2T}/T_0) &= P(S_{1F}, S_{2T}/T_0, B_{0T})P(B_{0T}) + P(S_{1F}, S_{2T}/T_0, B_{0F})P(B_{0F}) \\ &= P(S_{1F}, S_{2T}/T_0, B_{0T})P(B_{0T}) + 0 \\ &= P(S_{1F}/T_0, B_{0T})P(S_{2T}/T_0, B_{0T})P(B_{0T}) \\ &= (1 - f_n)f_n \cdot 0.9 \end{aligned}$$

- Denominator term,

$$\begin{aligned} P(S_{1F}, S_{2T}) = & P(S_{1F}, S_{2T}/T_0, B_{0T})P(T_0, B_{0T}) \\ & + P(S_{1F}, S_{2T}/T_0, B_{0F})P(T_0, B_{0F}) \\ & + P(S_{1F}, S_{2T}/T_{0F}, B_{0F})P(T_{0F}, B_{0F}) \\ & + P(S_{1F}, S_{2T}/T_{0F}, B_0)P(T_{0F}, B_0) \end{aligned}$$

the two terms in the middle are zero because if the battery is of it is not possible that the second sensor is true so this probability is zero,

$$\begin{aligned} & = P(S_{1F}, S_{2T}/T_0, B_{0T})P(T_0, B_{0T}) + P(S_{1F}, S_{2T}/T_{0F}, B_0)P(T_{0F}, B_0) \\ & = P(S_{1F}/T_0, B_{0T})P(S_{2T}/T_0, B_{0T})P(T_0)P(B_{0T}) + P(S_{1F}/T_{0F}, B_0)P(S_{2T}/T_{0F}, B_0)P(T_{0F})P(B_0) \\ & = (1 - f_n)f_n \cdot 0.2 \cdot 0.9 + f_p(1 - f_p) \cdot 0.8 \cdot 0.9 \end{aligned}$$

Finally we have that,

$$\begin{aligned} P(T_0 = T/S_1 = F, S_2 = T) & = \frac{P(S_1 = F, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = T)} \\ & = \frac{(1 - f_n)f_n \cdot 0.9}{(1 - f_n)f_n \cdot 0.2 \cdot 0.9 + f_p(1 - f_p) \cdot 0.8 \cdot 0.9} \cdot 0.2 \\ & = \frac{(1 - f_n)f_n}{(1 - f_n)f_n \cdot 0.2 + f_p(1 - f_p) \cdot 0.8} \cdot 0.2 \\ & \approx 0.42358 \end{aligned}$$

Case (a)

$$P(T_0 = T/S_1 = F, S_2 = F) = \frac{P(S_1 = F, S_2 = F/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = F)}$$

- Numerator term,

$$\begin{aligned} P(S_{1F}, S_{2F}/T_0) & = P(S_{1F}, S_{2F}/T_0, B_{0T})P(B_{0T}) + P(S_{1F}, S_{2F}/T_0, B_{0F})P(B_{0F}) \\ & = P(S_{iF}/T_0, B_{0T})^2 P(B_{0T}) + P(S_{iF}/T_0, B_{0F})^2 P(B_{0F}) \\ & = f_n^2 \cdot 0.9 + 0.1 \cdot 1^2 \end{aligned}$$

- Denominator term,

$$\begin{aligned} P(S_{1F}, S_{2T}) = & P(S_{iF}/T_0, B_{0T})^2 P(T_0)P(B_{0T}) \\ & + P(S_{iF}/T_0, B_{0F})^2 P(T_0)P(B_{0F}) \\ & + P(S_{iF}/T_{0F}, B_{0F})^2 P(T_{0F})P(B_{0F}) \\ & + P(S_{iF}/T_{0F}, B_{0T})^2 P(T_{0F})P(B_{0T}) \\ = & f_n^2 \cdot 0.2 \cdot 0.9 + 1^2 \cdot 0.2 \cdot 0.1 + 1^2 \cdot 0.8 \cdot 0.1 + (1 - f_p)^2 \cdot 0.8 \cdot 0.9 \\ \vdots & \\ = & 0.9 \left(0.2f_n^2 + 0.8(1 - f_p)^2\right) + 0.1 \end{aligned}$$

Finally we have that,

$$\begin{aligned}
P(T_0 = T/S_1 = F, S_2 = F) &= \frac{P(S_1 = F, S_2 = F/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = F)} \\
&= \frac{f_n^2 \cdot 0.9 + 0.1 \cdot 1^2}{0.9\,(0.2f_n^2 + 0.8(1 - f_p)^2) + 0.1} \cdot 0.2 \\
&\approx 0.02502004
\end{aligned}$$

Case (b)

We can now repeat the same procedure as in the previous case, so now we omit details and do it in less steps:

$$P(T_0 = T/S_1 = T, S_2 = T) = \frac{P(S_1 = T, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = T, S_2 = T)}$$

- Numerator term,

$$\begin{aligned}
P(S_{1T}, S_{2T}/T_0) &= P(S_{it}/T_0, B_{0T})^2 P(B_{0T}) + P(S_{it}/T_0, B_{0F})^2 P(B_{0F}) \\
&= P(S_{it}/T_0, B_{0T})^2 P(B_{0T}) + 0 \\
&= (1 - f_n)^2 \cdot 0.9
\end{aligned}$$

- Denominator term,

$$\begin{aligned}
P(S_{1T}, S_{2T}) &= P(S_{iF}/T_0, B_{0T})^2 P(T_0) P(B_{0T}) \\
&\quad + P(S_{iT}/T_{0F}, B_{0T})^2 P(T_{0F}) P(B_{0T}) \\
&= (1 - f_n)^2 \cdot 0.2 \cdot 0.9 + f_p^2 \cdot 0.8 \cdot 0.9
\end{aligned}$$

Finally we have that,

$$\begin{aligned}
P(T_0 = T/S_1 = T, S_2 = T) &= \frac{P(S_1 = T, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = T, S_2 = T)} \\
&= \frac{(1 - f_n)^2 \cdot 0.9}{(1 - f_n)^2 \cdot 0.2 \cdot 0.9 + f_p^2 \cdot 0.8 \cdot 0.9} \cdot 0.2 \\
&\approx 0.9995751
\end{aligned}$$

1.2 Kit B

Case (c)

$$P(T_0 = T/S_1 = F, S_2 = T) = \frac{P(S_1 = F, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = T)}$$

- Numerator term,

$$\begin{aligned}
P(S_{1F}, S_{2T}/T_0) &= P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2T})P(B_{1T}, B_{2T}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2F})P(B_{1T}, B_{2F}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2F})P(B_{1F}, B_{2F}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2T})P(B_{1F}, B_{2T}) \\
&= P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2T})P(B_{1T}, B_{2T}) \\
&+ 0 \\
&+ 0 \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2T})P(B_{1F}, B_{2T}) \\
&= P(S_{1F}/T_0, B_{1T}, B_{2T})P(S_{2T}/T_0, B_{1T}, B_{2T})P(B_{1T}, B_{2T}) \\
&+ P(S_{1F}/T_0, B_{1F}, B_{2T})P(S_{2T}/T_0, B_{1F}, B_{2T})P(B_{1F}, B_{2T}) \\
&= P(S_{1F}/T_0, B_{1T})P(S_{2T}/T_0, B_{2T})P(B_{1T}, B_{2T}) \\
&+ P(S_{1F}/T_0, B_{1F})P(S_{2T}/T_0, B_{2T})P(B_{1F}, B_{2T}) \\
&= f_n(1 - f_n) \cdot 0.9^2 + 1 \cdot (1 - f_n) \cdot 0.9 \cdot 0.1
\end{aligned}$$

- Denominator term,

$$\begin{aligned}
P(S_{1F}, S_{2T}) &= P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2T})P(T_0, B_{1T}, B_{2T}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2F})P(T_0, B_{1T}, B_{2F}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2F})P(T_0, B_{1F}, B_{2F}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2T})P(T_0, B_{1F}, B_{2T}) \\
&+ \text{the same four terms but with } T_{0F} \text{ instead of } T_0 \\
\text{non-zero terms} \rightarrow &= P(S_{1F}, S_{2T}/T_0, B_{1T}, B_{2T})P(T_0, B_{1T}, B_{2T}) \\
&+ P(S_{1F}, S_{2T}/T_0, B_{1F}, B_{2T})P(T_0, B_{1F}, B_{2T}) \\
&+ P(S_{1F}, S_{2T}/T_{0F}, B_{1T}, B_{2T})P(T_{0F}, B_{1T}, B_{2T}) \\
&+ P(S_{1F}, S_{2T}/T_{0F}, B_{1F}, B_{2T})P(T_{0F}, B_{1F}, B_{2T}) \\
&= P(S_{1F}/T_0, B_{1T})P(S_{2T}/T_0, B_{2T})P(T_0, B_{1T}, B_{2T}) \\
&+ P(S_{1F}/T_0, B_{1F})P(S_{2T}/T_0, B_{2T})P(T_0, B_{1F}, B_{2T}) \\
&+ P(S_{1F}/T_{0F}, B_{1T})P(S_{2T}/T_{0F}, B_{2T})P(T_{0F}, B_{1T}, B_{2T}) \\
&+ P(S_{1F}/T_{0F}, B_{1F})P(S_{2T}/T_{0F}, B_{2T})P(T_{0F}, B_{1F}, B_{2T}) \\
&= f_n(1 - f_n) \cdot 0.2 \cdot 0.9^2 \\
&+ 1 \cdot (1 - f_n) \cdot 0.2 \cdot 0.1 \cdot 0.9 \\
&+ (1 - f_p)f_p \cdot 0.8 \cdot 0.9^2 \\
&+ 1 \cdot f_p \cdot 0.8 \cdot 0.1 \cdot 0.9
\end{aligned}$$

Finally we have that,

$$P(T_0 = T/S_1 = F, S_2 = T) = \frac{P(S_1 = F, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = T)}$$
$$= \frac{f_n(1-f_n)\cdot 0.9^2 + 1\cdot(1-f_n)\cdot 0.9\cdot 0.1}{(f_n(1-f_n)\cdot 0.2\cdot 0.9^2 + 1\cdot(1-f_n)\cdot 0.2\cdot 0.1\cdot 0.9 + (1-f_p)f_p\cdot 0.8\cdot 0.9^2 + 1\cdot f_p\cdot 0.8\cdot 0.1\cdot 0.9}0.2$$
$$\approx 0.7565559$$

Case (a)

$$P(T_0 = T/S_1 = F, S_2 = F) = \frac{P(S_1 = F, S_2 = F/T_0 = T)P(T_0 = T)}{P(S_1 = F, S_2 = F)}$$

- Numerator term, (looking at the previous case we can deduce that only one term survives):

$$\begin{aligned} P(S_{1T}, S_{2T}/T_0) &= P(S_{1T}, S_{2T}/T_0, B_{1T}, B_{2T})P(B_{1T}, B_{2T}) \\ &= P(S_{iT}/T_0, B_{iT})^2 P(B_{1T}, B_{2T}) \end{aligned}$$

- Denominator term, only two terms survive,

$$\begin{aligned} P(S_{1T}, S_{2T}) = &P(S_{1T}, S_{2T}/T_0, B_{1T}, B_{2T})P(T_0, B_{1T}, B_{2T}) \\ &+ P(S_{1T}, S_{2T}/T_{0F}, B_{1T}, B_{2T})P(T_{0FF}, B_{1T}, B_{2T}) \end{aligned}$$

Finally we have that,

$$P(T_0 = T/S_1 = T, S_2 = T) = \frac{P(S_1 = T, S_2 = T/T_0 = T)P(T_0 = T)}{P(S_1 = T, S_2 = T)}$$
$$= \frac{P(S_{iT}/T_0, B_{iT})^2 P(B_{1T}, B_{2T})P(T_0)}{P(S_{1T}, S_{2T}/T_0, B_{1T}, B_{2T})P(T_0, B_{1T}, B_{2T}) + P(S_{1T}, S_{2T}/T_{0F}, B_{1T}, B_{2T})P(T_{0F}, B_{1T}, B_{2T})}$$
$$= \frac{P(S_{iT}/T_0, B_{iT})^2 P(T_0)}{P(S_{1T}, S_{2T}/T_0, B_{1T}, B_{2T})P(T_0) + P(S_{1T}, S_{2T}/T_{0F}, B_{1T}, B_{2T})P(T_{0F})}$$
$$= \frac{P(S_{iT}/T_0, B_{iT})P(T_0)}{P(S_{iT}/T_0, B_{iT})^2 P(T_0) + P(S_{iT}/T0F, B_{iT})^2 P(T_{0F})}$$
$$= \frac{(1-f_n)^2\cdot 0.2}{(1-f_n)^2\cdot 0.2 + f_p^2\cdot 0.8} = P_{kitA}(T_0/S_{1T}, S_{2T})$$
$$\approx 0.9995751$$

since it coincides with the case (b) we have computed for the kit A, by comparison.

Case (b)

This case is the one that have more terms to compute since all are non-zero. But it is straight forward and repetitive given the computations we have done so far.

2 Solutions using gRain

2.1 Kit A

With the following block of R-code we build the Bayesian Network for kit A, and from Figure 1a, in particular the tables of probabilities of each node.

```
tf <- c("false","true")
fp = 0.01*100
fn = 0.03*100

#####################
#     Kit A         #
#####################

  # Specify the CPTs:
  node.T0<- cptable(~ T0, values=c(8,2),levels=tf)
  node.B0<- cptable(~ B0, values=c(1,9), levels=tf)
  node.S1<- cptable(~ S1 + T0 + B0, values=c(1,0,1,0,100-fp,
   fp,fn,100-fn), levels=tf)
  node.S2<- cptable(~ S2 + T0 + B0, values=c(1,0,1,0,100-fp,
   fp,fn,100-fn), levels=tf)

  # Create an intermediate representation of the CPTs:
  nodesA <- list(node.T0,node.B0,node.S1,node.S2)
  plist <- compileCPT(nodesA)

  # Create a network of kit A:
  BNA<-grain(plist)
  summary(BNA)
```

In order to compute the probabilities of cases (a), (b) and (c) that we have already calculated “by hand” we consider the following code to do the queries:

```
  #################
  # Subexercise (a)#
  #################

    #Compute P(T0 = T/S1 = F, S2 = F)
    BNAa<-setEvidence(BNA,nodes=c("S1","S2"),
                     states=c("false", "false"))
    marga=querygrain(BNAa,nodes
                    =c("T0"), type="marginal")
    marga
    marga$T0["true"]
```

```
  #P(T0 = T/S1 = F, S2 = F) = 0.02502004

#################
# Subexercise (b)#
################
  #Compute P(T0 = T/S1 = T, S2 = T)
  BNAb<-setEvidence(BNA,nodes=c("S1","S2"),
                    states=c("true", "true"))
  margb=querygrain(BNAb,nodes
                   =c("T0"), type="marginal")
  margb
  margb$T0["true"]
  #P(T0 = T/S1 = T, S2 = T) = 0.9995751

#################
# Subexercise (c)#
################

  #Compute P(T0 = T/S1 = T, S2 = F)
  #Compute P(T0 = T/S1 = F, S2 = T)

  BNAc<-setEvidence(BNA,nodes=c("S1","S2"),
                    states=c("true", "false"))
  margc=querygrain(BNAc,nodes
                   =c("T0"), type="marginal")
  margc
  margc$T0["true"]
  #P(T0 = T/S1 = T, S2 = F) = 0.4235808
```

2.2 Kit B

With the following block of R-code we build the Bayesian Network for kit B, and from Figure 1b, in particular the tables of probabilities of each node.

```
######################
#     Kit B          #
######################
    # Specify the CPTs:
    node.T0<- cptable(~ T0, values=c(8,2),levels=tf)
    node.B1<- cptable(~ B1, values=c(1,9), levels=tf)
    node.B2<- cptable(~ B2, values=c(1,9), levels=tf)
    node.S1<- cptable(~ S1 + T0 + B1, values=c(1,0,1,0,100-fp
   ,fp,fn,100-fn), levels=tf)
    node.S2<- cptable(~ S2 + T0 + B2, values=c(1,0,1,0,100-fp
   ,fp,fn,100-fn), levels=tf)

    # Create an intermediate representation of the CPTs:
    nodesB <- list(node.T0,node.B1, node.B2,node.S1,node.S2)
    plistB <- compileCPT(nodesB)

    # Create a network of kit A:
    BNB<-grain(plistB)
    summary(BNB)
```

In order to compute the probabilities of cases (a), (b) and (c) that we have already calculated "by hand" we consider the following code to do the queries:

```
#################
    # Subexercise (a)#
    #################

    #Compute P(T0 = T/S1 = F, S2 = F)
    BNBa<-setEvidence(BNB,nodes=c("S1","S2"),
                      states=c("false", "false"))
    marga=querygrain(BNBa,nodes
                     =c("T0"), type="marginal")
    marga
    marga$T0["true"]
    #P(T0 = T/S1 = F, S2 = F)=0.004089033

    #################
    # Subexercise (b)#
    ################
    #Compute P(T0 = T/S1 = T, S2 = T)
```

```
BNBb<-setEvidence(BNB,nodes=c("S1","S2"),
                  states=c("true", "true"))
margb=querygrain(BNBb,nodes
                 =c("T0"), type="marginal")
margb
margb$T0["true"]
#P(T0 = T/S1 = T, S2 = T) = 0.9995751

#################
# Subexercise (c)#
################

#Compute P(T0 = T/S1 = T, S2 = F)
#Compute P(T0 = T/S1 = F, S2 = T)

BNBc<-setEvidence(BNB,nodes=c("S1","S2"),
                  states=c("true", "false"))
margc=querygrain(BNBc,nodes
                 =c("T0"), type="marginal")
margc
margc$T0["true"]
#P(T0 = T/S1 = T, S2 = F)  = 0.7565559
```

3 Discussion

"*In which of the three cases the answer is the same for the two kits? Why?* "

In the case when both sensors measure "true", i.e., when $S_1 = T$, $S_2 = T$ the probability of having extreme temperature, i.e, $T_0 = T$ is the same for both kits A and B, since the fact that the sensors are true determine that the batteries are on (otherwise, if they were off, then we know that the sensors would measure false with probability 1). Thus, in this case the role of the batteries is not influencing the probability of observing or not observing T_0, and that is why the results are the same for both kits and in this particular this case (b).

Mathematically speaking, we can look at the calculations by hand, and we will see that there occur some cancellations of terms that manifest what we have already said in a "more intuitive way" (in the precedent paragraph). ◀